

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-055444

(43)Date of publication of application : 24.02.1998

(51)Int.Cl. G06T 7/00
G06T 1/00

(21)Application number : 09-115461

(71)Applicant : LUCENT TECHNOL INC

(22)Date of filing : 06.05.1997

(72)Inventor : PODILCHUK CHRISTINE IRENE
ZHANG XIAOYU

(30)Priority

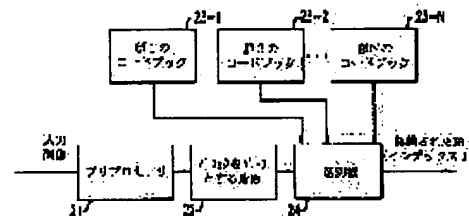
Priority number : 96 643685 Priority date : 06.05.1996 Priority country : US

(54) RECOGNITION OF FACE USING FEATURE VECTOR WITH DCT AS BASE

(57)Abstract:

PROBLEM TO BE SOLVED: To enable the recognition and identification of a human face especially included in an image by collating a prescribed object with one of known objects based on the error of coding.

SOLUTION: A preprocessor 21 extracts a concerned area (namely face) for an inputted image and standardizes the scale of reduction. The extracted section in the inputted image is divided into blocks and the respective blocks are converted with a clock as a base by transformation 22. When a DCT vector is once provided by the transformation 22, a sorter 24 is used for identifying the face included in the inputted face based on code books from 23-1 or 23-3 of the known face. For example, the sorter 24 judges which code books is most coincident with an input vector provided from the inputted image. In the case of face identification, the face related to the 'most coincident' code book is identified as the strongest candidate.



LEGAL STATUS

[Date of request for examination] 03.03.2000

[Date of sending the examiner's decision of rejection] 09.01.2002

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision 2002-06054
of rejection]

[Date of requesting appeal against examiner's 09.04.2002
decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19)日本国特許庁(JP)

(12)公開特許公報 (A)

(11)特許出願公開番号

特開平10-55444

(43)公開日 平成10年(1998)2月24日

(51)Int. Cl. ⁶	識別記号	庁内整理番号	FI	技術表示箇所
G 0 6 T	7/00		G 0 6 F	15/70 4 6 0 B
	1/00			15/62 3 8 0
				4 6 5 K

審査請求 未請求 請求項の数35 OL

(全12頁)

(21)出願番号 特願平9-115461

(22)出願日 平成9年(1997)5月6日

(31)優先権主張番号 08/643685

(32)優先日 1996年5月6日

(33)優先権主張国 米国 (US)

(71)出願人 596092698

ルーセント テクノロジーズ インコーポ
レーテッドアメリカ合衆国, 07974-0636 ニュージ
ャーシイ, マレイ ヒル, マウンテン ア
ヴェニュー 600(72)発明者 クリスチン アイレン ボデルチュック
アメリカ合衆国 08807 ニュージャーク
シイ, ブリッジウォーター, ビューモント
ウェイ 182

(74)代理人 弁理士 岡部 正夫 (外9名)

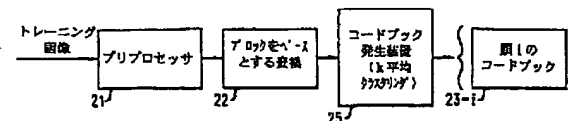
最終頁に続く

(54)【発明の名称】 DCTをベースとするフィーチャー・ベクトルを使った顔の認識

(57)【要約】

【課題】 DCTをベースとするフィーチャー・ベクトルを用いて顔の認識率を向上させる。

【解決手段】 入力画像信号中の所与のオブジェクトは、データベースに保存された複数の既知のオブジェクトの1つと比較され、その際、既知のオブジェクトの各々の保存された表示は、既知のオブジェクトを含むトレーニング画像信号に基づいて生成されたコードブックを含む。第1の例示としての実施例は、所与のオブジェクトをブロックに分解するステップと、各エンコーディングがコードブックの1つを使ったオブジェクトのコーディングからなる、所与のオブジェクトの複数のエンコーディングを行うステップと、各エンコーディングについてコーディング誤差を決定するステップと、コーディングの誤差に基づいて所与のオブジェクトを既知のオブジェクトの1つと比較するステップとを含む。



【特許請求の範囲】

【請求項 1】 入力画像信号に含まれる所与のオブジェクトを複数の既知のオブジェクトの 1 つと突き合わせる方法であって、既知のオブジェクトの各々の表示がデータベースに保存されており、それに対応するコードブックからなり、各コードブックが、対応する既知のオブジェクトからなる、1 つかそれ以上の対応するトレーニング画像信号に基づいて生成した複数のコードベクトルからなり、該方法が、

所与のオブジェクトを含む入力画像信号の一部を複数のブロックに分解するステップと、

各入力ベクトルが対応するブロックに基づく、複数のブロックに対応する複数の入力ベクトルを生成するステップと、

所与のオブジェクトを、1 つかそれ以上の入力ベクトルと、既知のオブジェクトに対応する 1 つかそれ以上のコードブックに含まれる 1 つかそれ以上のコードベクトルに基づく既知のオブジェクトの 1 つと突き合わせるステップを含む方法。

【請求項 2】 所与のオブジェクトを既知のオブジェクトの 1 つと突き合わせるステップが、

各エンコーディングが、対応する 1 つのコードブックに含まれる 1 つかそれ以上のコードベクトルに基づく 1 つかそれ以上の入力ベクトルのコーディングからなる、入力画像信号の少なくとも一部の複数のエンコーディングを行うステップと、

複数のエンコーディングに対応する複数のコーディング誤差を判断するステップと、

複数のコーディング誤差に基づいて、所与のオブジェクトを既知のオブジェクトの 1 つと突き合わせるステップを含む、請求項 1 に記載の方法。

【請求項 3】 コーディング誤差が平均二乗誤差からなる、請求項 2 に記載の方法。

【請求項 4】 複数のコーディング誤差を比較して、最も小さいコーディング誤差を決定するステップを含み、その際、所与のオブジェクトを既知のオブジェクトの 1 つと突き合わせるステップが、入力画像信号の少なくとも一部がエンコードされた時、最も小さいコーディング誤差を発生したコードブックに対応する既知のオブジェクトを選択するステップをさらに含む、請求項 2 に記載の方法。

【請求項 5】 所与のオブジェクトを既知のオブジェクトの 1 つと突き合わせるステップが、1 つかそれ以上の入力ベクトルに基づいて、所与のオブジェクトに対応するコードブックを生成するステップと、

所与のオブジェクトに対応するコードブックを、既知のオブジェクトに対応する 1 つかそれ以上のコードブックと比較するステップと、

所与のオブジェクトに対応するコードブックと、既知の

オブジェクトに対応するコードブックとの比較に基づいて、所与のオブジェクトを既知のオブジェクトの 1 つと突き合わせるステップを含む、請求項 1 に記載の方法。

【請求項 6】 所与のオブジェクトに対応するコードブックを既知のオブジェクトに対応するコードブックと比較するステップが、各比較について平均二乗誤差を計算するステップを含む、請求項 5 に記載の方法。

【請求項 7】 計算された平均二乗誤差から最小の平均二乗誤差を決定するステップをさらに含み、所与のオブジェクトを既知のオブジェクトの 1 つと突き合わせるステップが、所与のオブジェクトに対応するコードブックと比較する時最小の平均二乗誤差を生じるコードブックと対応する既知のオブジェクトを選択するステップを含む、請求項 6 に記載の方法。

【請求項 8】 入力ベクトルを生成するステップが、ブロックにリニア変換を行うステップと、変換されたブロックに基づいて入力ベクトルを生成するステップを含む、請求項 1 に記載の方法。

【請求項 9】 リニア変換が、ディスクリート・コサイン変換を含む、請求項 8 に記載の方法。

【請求項 10】 既知のオブジェクトに対応するコードブックがベクトル量子化コードブックを含む、請求項 1 に記載の方法。

【請求項 11】 既知のオブジェクトに対応するコードブックに含まれるコードベクトルが、k 平均クラスタリング技術を使って、1 つかそれ以上の対応するトレーニング画像信号から生成された、請求項 1 に記載の方法。

【請求項 12】 所与のオブジェクトを、所与のオブジェクトと一致する既知のオブジェクトとして識別するステップをさらに含む、請求項 1 に記載の方法。

【請求項 13】 所与のオブジェクトの要求されたアイデンティフィケーションが所与のオブジェクトと一致する既知のオブジェクトからなることを検証するステップをさらに含む、請求項 1 に記載の方法。

【請求項 14】 所与のオブジェクトと既知のオブジェクトの各々が人間の顔を含む、請求項 1 に記載の方法。

【請求項 15】 既知のオブジェクトの複数の表示を含むデータベースを生成する方法であって、データベースは入力画像信号に含まれる所与のオブジェクトを既知のオブジェクトの 1 つと突き合わせる際に使うためのものであって、既知のオブジェクトの表示の各々が対応するコードブックからなり、該方法が、

既知のオブジェクトの各々について、各トレーニング画像信号の一部が既知のオブジェクトを含む、1 つかそれ以上のトレーニング画像信号の各々の一部を複数のブロックに分解するステップと、

既知のオブジェクトの各々について、各トレーニング・ベクトルが対応するブロックに基づく、複数のブロックに対応する複数のトレーニング・ベクトルを生成するス

テップと、

既知のオブジェクトの各々について、既知のオブジェクトを含むトレーニング画像信号から生成された1つかそれ以上のトレーニング・ベクトルに基づいて、対応するコードブックを生成するステップとを含む方法。

【請求項16】 トレーニング・ベクトルを生成するステップが、ブロックにリニア変換を行うステップと、変換されたブロックに基づいてトレーニング・ベクトルを生成するステップとを含む、請求項15に記載の方法。

【請求項17】 リニア変換がディスクリート・コサイン変換を含む、請求項16に記載の方法。

【請求項18】 生成されたコードブックの各々がベクトル量子化コードブックを含む、請求項15に記載の方法。

【請求項19】 コードブックの各々が、k平均クラスタリング技術を使って、対応するトレーニング・ベクトルから生成する、請求項15に記載の方法。

【請求項20】 所与のオブジェクトと既知のオブジェクトの各々が人間の顔を含む、請求項15に記載の方法。

【請求項21】 入力画像信号からなる所与のオブジェクトを複数の既知のオブジェクトの1つと突き合わせるためのシステムであって、既知のオブジェクトの各々のためのデータベースに保存されて対応するコードブック表示がデータベースに保存されて対応するコードブックからなり、各コードブックが、対応する既知のオブジェクトを含む1つかそれ以上の対応するトレーニング画像信号に基づいて生成された複数のコードベクトルからなり、該システムが、

所与のオブジェクトを含む入力画像信号の一部を複数のブロックに分解するための手段と、

各入力ベクトルが対応するブロックに基づく、複数のブロックに対応する複数の入力ベクトルを生成するための手段と、

1つかそれ以上の入力ベクトルと、既知のオブジェクトに対応する1つかそれ以上のコードブックに含まれる1つかそれ以上のコードベクトルに基づいて、所与のオブジェクトを既知のオブジェクトの1つと突き合わせるための手段とを備えてなるシステム。

【請求項22】 所与のオブジェクトを既知のオブジェクトの1つと突き合わせるための手段が、

各エンコーディングが、コードブックの対応する1つに含まれる1つかそれ以上のコードベクトルに基づく1つかそれ以上の入力ベクトルのコーディングを含む、入力画像信号の少なくとも一部分の複数のエンコーディングを行うための手段と、

複数のエンコーディングに対応する複数のコーディング誤差を判断するための手段と、

複数のコーディング誤差に基づいて、所与のオブジェクトを既知のオブジェクトの1つと突き合わせるための手段とを備えてなる、請求項21に記載のシステム。

【請求項23】 複数のコーディング誤差を比較して最小のコーディング誤差を決定するコンパレータをさらに含み、所与のオブジェクトを既知のオブジェクトの1つと突き合わせるための手段が、入力画像信号の少なくとも一部がエンコードされた時、最も小さなコーディング誤差を発生したコードブックに対応する既知のオブジェクトを選択するセレクトを含む、請求項22に記載の方法。

【請求項24】 所与のオブジェクトを既知のオブジェクトの1つと突き合わせるための手段が、1つかそれ以上の入力ベクトルに基づいて、所与のオブジェクトに対応するコードブックを生成するための手段と、

所与のオブジェクトに対応するコードブックを、既知のオブジェクトに対応する1つかそれ以上のコードブックと比較するコンパレータと、

所与のオブジェクトに対応するコードブックと、既知のオブジェクトに対応するコードブックとの比較に基づいて、所与のオブジェクトを既知のオブジェクトの1つと突き合わせるための手段とを備えてなる、請求項21に記載のシステム。

【請求項25】 コンパレータが各比較について、平均二乗誤差を計算する、請求項24に記載のシステム。

【請求項26】 計算された平均二乗誤差から最小の平均二乗誤差を決定するための手段をさらに含み、所与のオブジェクトを既知のオブジェクトの1つと突き合わせるための手段が、既知のオブジェクトと対応するコードブックと比較される時最小の平均二乗誤差となるコードブックに対応する既知のオブジェクトを選択するセレクトを含む、請求項25に記載のシステム。

【請求項27】 所与のオブジェクトに対応するコードブックを生成するための手段が、k平均クラスタリング技術を行うための手段を含む、請求項24に記載のシステム。

【請求項28】 入力ベクトルを生成するための手段が、ブロックにリニア変換を行うための手段と、変換されたブロックに基づいて入力ベクトルを生成するための手段とを備えてなる、請求項21に記載のシステム。

【請求項29】 所与のオブジェクトを、所与のオブジェクトと一致する既知のオブジェクトとして識別するための手段をさらに含む、請求項21に記載のシステム。

【請求項30】 所与のオブジェクトの要求されたアイデンティフィケーションが、所与のオブジェクトと一致した既知のオブジェクトを含むことを検証するための手段をさらに含む、請求項21に記載のシステム。

【請求項31】 所与のオブジェクトと、各既知のオブジェクトが人間の顔を含む、請求項21に記載のシステム。

【請求項32】 既知のオブジェクトの複数の表示からなるデータベースを生成するためのシステムであって、

そのデータベースが入力画像信号に含まれる所与のオブジェクトを既知のオブジェクトの1つと突き合わせる際に使うためのもので、既知のオブジェクトの表示の各々が対応するコードブックを含み、該システムが、既知のオブジェクトの各々に適用される、各トレーニング画像信号の一部が既知のオブジェクトを含む、1つかそれ以上のトレーニング画像信号の各々の一部を複数のブロックに分解するための手段と、

既知のオブジェクトの各々に適用される、各トレーニング・ベクトルが対応するブロックに基づく、複数のブロックに対応する複数のトレーニング・ベクトルを生成するための手段と、

既知のオブジェクトの各々に適用される、既知のオブジェクトを含むトレーニング画像信号から生成された1つかそれ以上のトレーニング・ベクトルに基づいて、対応するコードブックを生成するための手段とを備えてなるシステム。

【請求項 33】 トレーニング・ベクトルを生成するための手段が、ブロックにリニア変換を行うための手段と、変換されたブロックに基づいてトレーニング・ベクトルを生成するための手段とを備えてなる、請求項 32 に記載のシステム。

【請求項 34】 コードブックを生成するための手段が、k平均クラスタリング技術を行うための手段を備えてなる、請求項 32 に記載のシステム。

【請求項 35】 所与のオブジェクトと各既知のオブジェクトが人間の顔を含む、請求項 32 に記載のシステム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、概して、自動化画像分析の分野に関し、特に画像に含まれる人間の顔を認識および識別する問題に関する。

【0002】

【従来の技術、及び、発明が解決しようとする課題】検証と識別を含む、静止画像またはビデオ・データからの顔の認識は大きなデータベースのビデオ検索、保安上のアクセスなどの無数のマルチメディア適用例を含む広い範囲の商業的な適用例がある。顔を認識する問題は、一般に以下のように説明される。静止画像またはビデオ・シーケンスが入力として与えられると、必要とされるタスクは、例えば以前に保存された顔のデータベースの内容に基づいて、入力された画像の中に1つかそれ以上の人物を認識することである。高品質の顔認識システムは、照明、背景、顔の表情、位置、縮尺に変化がある時や、入力された画像とデータベースの元になるトレーニング画像の間に髪型の違いがあっても、入力された画像すなわちテスト画像と保存されたデータベースを一致させることができるはずである。

【0003】顔識別の問題は、詳細には、入力された画

像の人物が保存されたデータベースに含まれる人物の誰かに一致するか、もし一致するなら、どの人物と一致するかを判定する、一対多マッピングをシステムが行うことを要求する。他方、顔の認証の問題は、入力された画像が(十分に)データベースの中の1人の特定の個人のデータ(通常そのアイデンティティがシステムを使う人物によって要求される個人のためのデータ)に一致するかを判定する、一対一マッピングをシステムが行うことを要求する。顔の認証がセキュリティの分野で利用される一方、顔の識別はセキュリティと同様、新聞報道のような分野で利用され、そこでは多数のマルチメディア・データベースから特定の個人の画像を検索できることが望ましい。

【0004】通常の顔認識システムは、3つの部分、すなわち、データを分割して重要な部分または特徴を抽出する前処理ステップ、特徴選択ステップ、分類付けステップからなる。例えば、前処理ステップは、入力された画像を粗分割して背景のデータから顔のデータを分離するステップからなるか、または特徴を生じるために使われる目、鼻、口といった顔の部分の位置を見つけるためのより細かな分割を含む。前処理ステップはまた、顔の部分同士の位置関係といったフィーチャー・ベクトルを抽出する。いずれにしても、一度前処理がなされると、対象となる部分、すなわち顔は、通常画像の他の部分から抽出されて(寸法と位置について)標準化される。こうした前処理は従来通りで、本技術に普通に熟練した者には周知であるので、ここでの残りの議論は顔認識プロセスの特徴抽出および分類付け部分に焦点を当てる。

【0005】重要な顔の部分抽出したり、または特定の顔の部分同士の位置関係を判定して、それらをデータベースと比較する識別用の特徴として使うことに基づく多数の顔認識のための技術が使われたり、および/または提案されてきた。あいにく、これらの方法は通常顔の部分非常に高い精度で位置決めする能力に依存する。このことは画像が非常に良く管理された環境で捉えられないと、非常に困難である。さらに、これらの技術の大部分は画像のエッジの正確な検出や、照明の変化や多くの量のノイズを伴う質の低い画像に対して特に強くないローカル操作に依存する。

【0006】例えば、眼の間の距離、口の幅、眼から鼻への距離、眼から口までの距離といった位置関係を測定し、これらのパラメータを、それに対して入力(すなわち、テスト)データとトレーニング・データ(すなわち、データベースに保存されたデータ)の間で突き合わされる特徴として使おうとする技術は、例えば、R. Brunelli, T. Poggio, 「位置的特徴を通じた顔の認識」、Proc. Europe. Conf. on Computer Vision, pp. 792-800, 1992を含む。

【0007】他方、特徴選択プロセスのための顔の部分

の位置決めに依存しない技術は、例えば、顔の認識のために良く知られたKarhunen-Loeve変換を使うL. Sirovich, M. Kirby, 「人間の顔の特徴づけのためのロー・ディメンショナル手続き」、J. Opt. Soc. Amer., vol 4, p p. 519-524と、顔の検出と認識のために「固有の顔」を定義して使用するM. A. Turk, A. P. Pentland, 「固有の顔を使った顔の認識」Proc. Int. Conf. on Patt. Recogn., pp. 586-591, 1991を含む。これらのアプローチの各々では、各人の顔は本質的に、あらかじめ定義された「固有の顔」の集合に適用される重みのベクトルとして示される。この特定のアプローチの弱点の1つは、入力された画像が「固有の顔」を計算するために収集された画像と縮尺や頭の位置の点で同じでない場合、認識率が大きく低下する点である。

【0008】最後に、顔の認識のためにニューラル・ネットワークも利用される。しかし、ニューラル・ネットワークのアプローチは、通常大量のトレーニング・データ（例えば、データベース中の各人について、数百の画像）を必要とするので、大きなデータベースが必要になるため、このアプローチを非実用的なものにしている。上記で説明された従来技術のアプローチは他のものと同様、R. Chellappa, C. L. Wilson, S. Sirohey, 「人間と機械による顔の認識：概観」、Proceedings of the IEEE, 1995年5月で説明されている。

【0009】

【課題を解決するための手段】本発明は、顔の画像データに固有の特性の何らかの事前知識を組み込むことなく、一般的な画像の統計を利用する信号処理アプローチを利用するので有利である。もちろん、これは、大部分が発見的方法に基づき、一般的に、広い範囲の入力画像の変化に対して特に強くない、上記で説明された従来技術の「コンピュータ・ビジョン」アプローチとは対照的である。詳細には、本発明によれば、画像の中の顔というオブジェクトの認識は、認識されるオブジェクト（例えば、顔）を含む画像の一部がコード化され、それによって生じたコード化データが、既知のオブジェクト・データベースに保存された、前にコード化されたトレーニング・データと一致する画像の一部を含む、信号圧縮技術（すなわち、コーディング技術）を使って達成される。画像圧縮技術を使うことによって、各オブジェクト（例えば、人物の顔）は、小さな数の集合によって表すことができるので有利である。

【0010】例えば、本発明のある例示としての実施例によれば、オブジェクト（例えば、顔）を含む入力画像の一部がピクセル・データのブロックに分解され、各ブロックのピクセル・データに対して、例えば、ディスク

リット・コサイン変換(DCT)のような(リニア)変換が行われ、変換されたブロックは「ベクトル量子化」(VQ)され、各人物についてコードベクトルのコードブックを発生する。(ブロックをベースとする変換はブロックの信号エネルギーを、少数の有意な変換係数にコンパクト化する効果を持ち、この係数は有利にもブロックを表すために使用できるので有利である。トレーニング集合の中のデータの各ブロックを示す変換係数は、コードベクトルのコードブックを生成するために使われる。)一般に、入力されるベクトルの数は、コードブックに保存されたコードベクトルの数よりはるかに大きい。従って、トレーニング・データを示すために必要なデータの量はさらに圧縮されるので有利である。

【0011】ブロックをベースとする変換(DCTを含む)とベクトル量子化(VQ)は共に画像コーディングの分野で共通して使われる周知の技術であり、本技術に普通に熟練した者には周知の技術である。ディスクリット・コサイン変換は、例えばN. Ahmed, T. Natarajan, K. Rao, 「ディスクリット・コサイン変換」、IEEE Trans. Comput., vol 23, pp. 90-93, 1974に説明され、ベクトル量子化は、例えばY. Linde, A. Buzo, R. M. Gray, 「ベクトル量子化設計のためのアルゴリズム」、IEEE Trans. Commun., COM-28, pp. 84-95, 1990に説明されている。トレーニング・データからコードブックのコードベクトルを生成するために有利にも利用される例示としての技術の1つは「k平均クラスタリング」として知られ、これも本技術に普通に熟練した者には周知の技術で、例えば、S. P. Lloyd, 「PCMにおける最小二乗量子化」、IEEE Trans. Inform. Theory, IT-28, pp. 127-135, 1982に説明されている。「ディスクリット・コサイン変換」、「ベクトル量子化設計のためのアルゴリズム」、「PCMにおける最小二乗量子化」はそれぞれ、ここで発表されたのと同様に、参照によってここに組み込まれている。

【0012】本発明のある例示された実施例によれば、入力画像信号の中の所与のオブジェクトはデータベースの中に保存された複数の既知のオブジェクトの1つと一致するが、その際、各既知のオブジェクトの保存された表示は、既知のオブジェクトを含むトレーニング画像信号に基づいて生成するコードブックを含む。第1の例示としての実施例は、所与のオブジェクトをブロックに分解するステップと、各エンコーディングがコードブックの1つを使ったオブジェクトのコーディングからなる、所与のオブジェクトの複数のエンコーディングを行うステップと、各エンコーディングについてコーディングの誤差を判定するステップと、コーディングの誤差に基づいて、所与のオブジェクトを既知のオブジェクトの1つと突き合わせるステップとを含む。第2の例示としての

実施例は、所与のオブジェクトをブロックに分解するステップと、ブロックに基づいて所与のオブジェクトに対応するコードブックを生成するステップと、所与のオブジェクトに対応するコードブックを各既知のオブジェクトに対応するコードブックと比較するステップと、所与のオブジェクトをコードブックの比較に基づいて既知のオブジェクトの1つと突き合わせるステップとを含む。

【0013】

【発明の実施の形態】図1は、既知の顔のためのデータが、本発明の例示としての実施例に従って構成されるデータベース発生プロセスの略図である。各々が既知の個人の顔の異なった「絵」を含む多数の異なったトレーニング画像が、有利にも図1のプロセスに提供され、それによって、異なった照明、背景、顔の表情、姿勢、縮尺等といった種々の条件下での所与の既知の顔の画像をより完全に反映するコードブックが作り出される。詳細には、図1のプロセスは有利にも、その顔がデータベースに含まれている各人物について繰り返されるので、異なったコードブックが、トレーニング集合の中の異なった人物の各々について、図1のプロセスによって生成される。図示されるように、1つかそれ以上のトレーニング画像（各々が、トレーニング集合の中のi番目の人物の顔の画像を含む）が処理され、その1人の特定の人物のためのコードベクトルの集合を含む、コードブック23-iとなる。

【0014】図面を参照すると、（上記で説明したように）第1プリプロセッサ21は入力された画像から関心のある領域（すなわち、顔）を抽出し、縮尺を標準化する。従って、それ以降の処理はすべて、有利にも顔について（しかも顔だけについて）行うことができ、システムに提出される顔はすべて同じ縮尺（すなわち、サイズ）となる。プリプロセッサ21によって行われる前処理は従来通りで本技術に熟練した者には周知である。

【0015】前処理が完了した後、画像の抽出された部分はブロックに分解され、各ブロックに対して、変換22によって、リニアな、ブロックをベースとする変換が行われる。ブロックは、例えば、 8×8 の寸法の領域（すなわち、8ピクセル対8ピクセルの寸法の正方形）からなるか、または、 16×16 の寸法でも良い。ブロックは画像の範囲にわたって隣接しているか、または、好適には、重なっている（例えば、各範囲について2ピクセルずつ）。リニア変換の演算子を元の信号に適用することによって、より有効な表現が得られ、そこではデータは統計的に独立（または少なくとも非相関的）である。さらに、リニア変換はまた、一般にすべての信号圧縮技術にとって望ましい特性である、大きなエネルギーのコンパクト化を提供する。画像データ圧縮の分野で普通に使われる変換の1つが2方向ディスクリート・コサイン変換（DCT）であり、本技術に普通に熟練した者には周知である。DCTは上記で言及した長所の各々を

提供し、現在静止画像と動画の両方について、すべての国際圧縮規格のための基礎単位である。

【0016】変換22の変換を各画像ブロックに行うことによって、これらのブロックの各々は少数のDCT係数によって表示できるので有利である。特に、変換22は、 $(B \times B)$ ブロックの変換から得られる $(B \times B)$ 係数から、最大の信号エネルギー内容を持つ（L）係数のみを提供する。例えば、ブロックの寸法が 8×8 （すなわち、 $B=8$ ）の時、 $L=8$ というDCTの係数が提供されるので有利である。また、例えば、ブロックの寸法が 16×16 （すなわち、 $B=16$ ）の時、 $L=16$ というDCTの係数が提供されるので有利である。本技術に普通に熟練した者には周知であるように、最大の信号エネルギー内容を持つDCT係数は通常低い周波数の係数に対応する。

【0017】トレーニング・データ中の各ブロックについて（生成され提供された）DCT係数は、組み合わされて対応するベクトルとなり、その結果生じたベクトルの集合は、コードブック発生装置25によって、VQコードブックをデザインするために使われる。このコードブックはデータベース中の所与の既知の個人の顔を表すデータを含んでいる。コードブック・ベクトルの数は、通常それらが生成される元となった（DCT）ベクトルの総数よりはるかに小さいことに留意されたい。例えば、1000以上の入力ベクトルが有利にもコードブック発生装置25に供給され、それらは100以下（例えば64）のコードブック・ベクトル（すなわち、コードベクトル）を含むコードブックにマッピングされる。コードブック発生装置25によって行われるコードブックのデザインは、例えば、良く知られた「k平均クラスタリング」技術か、または信号圧縮の分野で良く知られている、トレーニング・ベクトルのデータからコードブックをデザインするための従来の技術からなる。

【0018】図3は図1の例示としてのトレーニングの手続きのコードブック発生装置25によって行われるコードブック発生プロセスの例示としての流れ図である。上記で説明したように、有利にも顔がデータベースに示される各人物について行われる、図1のプロセスの終了結果は、その人物の顔を含む1つかそれ以上のトレーニング画像に基づいて、所与の人物のためのコードブックを生成することである。従って、コードブック C^p が、各人物について生成されるので、 $p=1, 2, \dots, N$ 。この時Nは、データベース中の人物（すなわち、顔）の数である。各コードブックは複数のコードベクトル y_i からなり、この時 $i=1, 2, \dots, M$ であり、この時Mは各コードブックのサイズ（すなわち、そこに含まれるコードベクトルの数）である。この例では、Mは64に等しい。各コードベクトルは、上記で説明したように、Lの値（すなわち、DCT係数）からなる。この例では、ブロックは 8×8 の寸法であり、Lは8であ

る。

【0019】図3に示す例示としての手続きは、トレーニング・ベクトルを一連の「中間」コードブック C^p_0 、 C^p_1 、...、 C^p_m ...と繰り返し突き合わせ、各反復の際に、最終的な基準が合致するまでコードブックを変更する(すなわち、コードブック C^p_m を改良されたコードブック C^p_{m+1} によって置き換える)、反復プロセスを含む。最終的な反復から生じたコードブックは、データベースの中の特定の個人の顔を表すために使われるコードブックになるので有利である。

【0020】詳細には、図3の例示としての手続きは、最初の反復で使われるコードブック C^p_0 を初期化して、コードベクトルの無作為の集合にする、ステップ31から始まる。これらの無作為に選ばれたコードベクトルは、例えば、トレーニング集合からの最初のMベクトルを含むか、または、無作為の値を生成することによって選択される。その後、各反復mで、ステップ32は、トレーニング・ベクトルの(完全な)集合を使って、「クラスタ集合」を形成するが、それは本技術に熟練した者には周知である。詳細には、各トレーニング・ベクトル x_j (この時 $j=1, 2, \dots, J$)は、コードブック C^p_m の中の各コードベクトル y^p_i (この時 $i=1, 2, \dots, M$)と突き合わされ、誤差関数 $d(x_j, y^p_i)$ を計算する。この突き合わせの結果、各トレーニング・ベクトル x_j は、対応する1つのコードベクトル y^p_i に関連するクラスタの集合 R_i に割り当てられ、すなわち以下の式となる。

$R_i = \{x: d(x, y^p_i) < d(x, y^p_k) \text{ すべてに } k \neq i\}$

すなわち、各クラスタの集合 R_i は、すべてのトレーニング・ベクトル x からなり、トレーニング・ベクトル x とi番目のコードベクトル y^p_i の間の誤差(すなわち、差)は、トレーニング・ベクトル x とコードブックの中の他のどのコードベクトルとの間の誤差よりも小さい。言い換えれば、各トレーニング・ベクトルは、誤差関数 $d(x, y)$ によって決定される、「最も近い」現行のコードブックの中のコードベクトルに関連するクラスタに割り当てられる。誤差関数 $d(x, y)$ は何らかの適当な誤差の計量である。例えば、 $d(x, y)$ は、本技術に普通に熟練した者には周知の、従来の平均二乗誤差(mse)の計量である。

【0021】ステップ33では、新しいコードブック C^p_{m+1} が、ステップ32の分析から生じたクラスタの集合に基づく次の反復で使うために展開される。詳細には、各クラスタの集合 R_i について、集合の重心(すなわち、集合の中の独立したベクトルのn次元の加重平均)が計算される。その後、(次の反復のための)新しいコードブックが、各 $i=1, 2, \dots, M$ について、計算された重心 $\text{cent}(R_i)$ を含むように構成される。誤差の計量として平均二乗誤差が使われる場合、重

心は減少して算術的平均になることに留意されたい。すなわち、 $R_i = x_j$ ($j=1, 2, \dots, \|R_i\|$)であり、 $\|R_i\|$ が集合 R_i の濃度(すなわち、元の個数)である時、次の式になる。

【数1】

$$\text{cent}(R_i) = \frac{1}{\|R_i\|} \sum_{j=1}^{\|R_i\|} x_j$$

10 【0022】ステップ34では、所与の反復に関する総ひずみの大きさ D_m が、トレーニング・ベクトルの各々とそれがクラスタをなす重心の間の距離に基づいて計算される。すなわち、

$$D_m = \sum d(x_j, \text{cent}(R_i))$$

であり、この時、合計はすべての $j=1, 2, \dots, J$ について(すなわち、すべてのトレーニング・ベクトルについて)行われる。こうして、 D_m は、コードブック C^p_m がトレーニング・ベクトルの所与の集合をコード化する正確さの度合いを含む。最後に、決定ステップ35は、前の反復から現在の反復へのコードブックの質の改善の度合いを測定するために、計算されたひずみの大きさ(D_m)と前の反復の際計算されたひずみの大きさ(D_{m-1})の間の差を計算する。この差が所定のしきい値 T_0 以下である場合、さらに反復を行うことによって得られそうなコードブックの質の改善は小さいので、反復は有利にも終了される。

【0023】一度既知の顔を表すコードブックのデータベースが作り出されると、顔の認識が、本発明の原理に従って、データベースに含まれるコードブックの集合に基づいて行われる。図2は、本発明の例示としての実施例による顔認識プロセスの略図である。例えば、図1および図3の例示としての手続きを使って生成したものと、コードブックのデータベース(例えば、コードブック23-1, 23-2, ..., 23-N)を与えられ、図2に示すような例示としての実施例は、入力画像信号からなる人物の顔を識別する(または顔の同一性を検証する)ために使われる。

【0024】第1に、図1の例示としてのコードブック発生プロセスと同一の方法で、プリプロセッサ21は入力された画像から関心のある領域(すなわち、顔)を抽出し、縮尺を標準化する。従って、これ以降の処理はすべて、有利にも顔について、(しかも顔だけについて)行われ、システムに提示される顔はすべて同じ縮尺(すなわち、サイズ)である。ここでも、プリプロセッサ21によって行われる前処理は、従来通りであり、本技術に熟練した者には周知である。

【0025】前処理が完了した後、再び図1の例示されたコードブック発生プロセスと同一の方法によって、入力された画像の抽出された部分がブロックに分解され、変換22によって、リニアな、ブロックをベースとする

変換が各ブロックについて行われる。ここでも、ブロックは、例えば、 8×8 の寸法かまたは 16×16 の寸法であり、それらは画像空間にわたって隣接しているか、好適には、重なり合っている。コードブック23-1から23-Nの生成の際に使われたのと同じ(DCTのような)変換が、図2の顔認識プロセスでも使われるので有利である。さらに、同じブロックのサイズが使われるので有利であり、同じ(L)係数の集合が提供されるので有利である。

【0026】一度DCT入力ベクトルが変換22によって提供されると、選別機24が、(既知の)顔のコードブック23-1から23-Nに基づいて、入力された画像に含まれる顔を識別する(または、顔の検証の場合、検証する)ために使われる。詳細には、選別機24は、例えば、どのコードブックが入力された画像から得られた入力ベクトルと最も良く一致するかを判断する。顔識別の場合、「最も一致する」コードブックに関連する顔が最も有力な候補として識別される。しかし、比較の基礎となる誤差の計量(例えば、以下の議論を参照)が、既知の顔のコードブックのいずれについても「十分に」小さな総誤差の値(例えば、所定のしきい値以下)を提供せず、(それによって「十分に良い」組み合わせが見つからないことが示される)場合、システムは既知の顔のいずれをも入力された画像の顔として認識しない。むしろ、システムは単純に、顔が画像の中にある人物は、実際に、顔のデータが以前に収集され、データベースに保存された人々の中の誰でもないと判断する。

【0027】顔の検証の場合、システムは、要求されたアイデンティティが、入力された画像の中の顔と最も良く一致すると判断された既知の顔のコードブックと一致するかどうかに基づいて、要求されたアイデンティティを検証または拒絶(すなわち、検証に失敗)する。または、システムは入力された画像が、要求されたアイデンティティの顔に関連する特定のコードブックに関して分析された時得られる総誤差の計量が「十分に」小さく、それによって入力された画像の中の顔と要求されたアイデンティティの顔の間に十分「良好な」一致が示されるかどうかに基づいて、要求されたアイデンティティを検証または拒絶する。

【0028】本発明のある実施例によれば、どの既知の顔のコードブックが入力された画像の顔と最良の一致を提供するかの決定は、多数の異なった方法で行うことができる。例えば、図2の例示としての顔認識プロセスの選別機24の動作の2つの例示としての実施例が、それぞれ図4および図5に示される。

【0029】図4は、本発明の第1の例示としての実施例による図2の例示としての顔認識プロセスの選別機24の動作の流れ図である。この実施例によれば、既知の顔のコードブックの各々は入力画像ベクトルをエンコーディングするために使われ、総コーディング誤差はその

エンコーディングの各々について決定される。その後、総コーディング誤差の最も低いエンコーディングが決定され、そのエンコーディングで使われたコードブックに対応する顔が、入力された画像の顔と最も良く一致する顔として識別される。

【0030】詳細には、図4を参照すると、例示としてのプロセスはステップ41から始まり、そこでは入力された画像ベクトルの複数の従来のVQコードブックのエンコーディングが行われ、こうしたエンコーディングの各々は既知の顔のコードブック(すなわち、コードブック23-1から23-N)の異なった1つを使って行われる。詳細には、また本技術に普通に熟練した者には周知であるように、これらのエンコーディングの各々は、入力された画像ベクトルのための所与のコードブックからのコードベクトルの1つを選択することからなり、その際、選択は、例えば平均二乗誤差のような誤差の計量に基づいてなされる。

【0031】選択されたコードベクトルは、入力されたベクトルとの「最良の一致」を提供するものである。すなわち、コードベクトルは、各入力ベクトルについて、入力ベクトルとコーディング(すなわち、選択されたコードベクトル)の間の誤差(すなわち、差)を有利にも最小にするように選択される。従って、コードベクトル y^{p_i} を含む各コードブック C^p (この時 $p=1, 2, \dots, N$)について、各入力ベクトル x_j (この時 $j=1, 2, \dots, J$)は各コードベクトルと比較され、最小の誤差を生じる1つが決定され、誤差関数 $d(x, y)$ が与えられる。詳細には、コードブック C^p を使った入力ベクトル x_j のコーディングに関する最小誤差を表す $d_{min}^p(x_j)$ は、各入力ベクトル x_j について、以下のように決定される。

【0032】 $d_{min}^p(x_j) = \{d(x_j, y^{p_i}) : d(x_j, y^{p_i}) < d(x_j, y^{p_k})$

すべて $k \neq i$, 但し $i, k=1, 2, \dots, M\}$

【0033】各エンコーディングが行われた後、図4の例示としての手続きのステップ44は、各エンコーディングについて(すなわち、すべての $p=1, 2, \dots, N$ について)総誤差 D^p を計算する。これらの総誤差は単に個別のコーディング誤差の合計であり、以下のように計算される。

$D^p = \sum d_{min}^p(x_j)$

この時合計はすべての $j=1, 2, \dots, J$ について行われ、 $d_{min}^p(x_j)$ の個別の値は、上記で示したように、各入力ベクトル x_j のコーディングの際に、対応するコードブック C^p を使って決定される。従って、これらの総誤差の値はコードブックの各々について行われたエンコーディングの全体的な質(すなわち、正確さ)を表す。(もちろん、ステップ44が、ステップ41で行われるエンコーディングから分離した別個の動作として行われる必要を回避して、これらの総誤差の値がま

10

20

30

40

50

た、有利にもエンコーディング・プロセスの間にそれ自体の一部として計算されることもあることに留意されたい。)

【0034】図4の例示としての手続きを完了すると、ステップ45は、エンコーディング全体から得られた最小の総誤差の値(すなわち、 D^p が最小になる p の値)を判断し、ステップ46は、最小の総誤差(すなわち、最上のエンコーディング)を達成したエンコーディングの際に使われたコードブックによって示される既知の顔を選択する。この方法で、入力された画像から得られた顔は、関連するコードブックが与えられた顔の「最良の」エンコーディングを生じるデータベースの中の既知の顔として識別されるので有利である。

【0035】図5は、本発明の第2の例示としての実施例による図2の例示としての顔認識プロセスの選別機24の動作の流れ図である。この例示としての実施例は、図4の例示としての実施例よりもわずかに性能を低下させるだけで、コンピュータの使用上より簡単な技術が提供される。詳細には、この第2の例示としての実施例は、入力された画像についてコードブックを生成し、その後「最上の一致」を決定するためにこのコードブックを既知の顔のコードブックの各々と比較する。ここでも、図4の例示された手続きのように、最も低い総誤差が、最上の一致を示すものと考えられる。しかし、図4および図5の例示としての手続きの間の差は、(J)個の入力ベクトルの各々のコーディング(それは、例えば、数千のベクトルのコーディングを必要とする)よりもむしろ、図5の手続きの中で、ステップ42が例示するように、 M 個(例えば64)のコードベクトルだけを含むコードブックを発生することにある。その後、このより制限されたベクトルの集合(すなわち、通常 $J \gg M$ の時、 J 個の入力ベクトルでなく、 M 個のコードベクトル)が、 N 個のコードブックを使ってエンコードされ(ステップ43)、これらの(簡単な)エンコーディングの各々から得られた総誤差が(ステップ44~46で)使われて、最小の総誤差を持つコードブックによって表される既知の顔を選択する。

【0036】詳細には、ステップ42のコードブックの生成は、本質的には図3で示したような、既知の顔のコードブックを生成する単に例示的に使われたのと同じ手続きからなる。しかし、この場合、図3の例示としてのトレーニングの手続きは、トレーニング画像データではなく、入力画像データに適用される。(トレーニング画像ベクトルが有利にも得られる潜在的に多くの画像があるのに対して、入力画像ベクトルが得られるのは通常1つの画像だけしかないことに留意すべきである。)従っ

て、図4の例示としての手続きの代わりに、図5の例示としての手続きを使うことの利点は、 N 個のコードブックの各々について、 J 個の(入力画像から得られた)コードベクトルを、 M 個の(トレーニング画像から得られた)コードベクトルと比較しなければならないのとは違って、 N 個のコードブックの各々について、 M 個の(入力画像から得られた)コードベクトルを M 個の(トレーニング画像から得られた)コードベクトルと比較するだけで良いということである。通常、 $J \gg M$ なので、このことは計算量の大きな節約になるので有利である。ここでも、使われる誤差の計量は平均二乗誤差であるか、または、本技術に普通に熟練した者には周知の、何らかの他の誤差の計量でもよい。

【0037】ここでは本発明の多数の特定の実施例が示され説明されたが、これらの実施例は、本発明の原理の適用例の中で考案できる多くのありうる特定の配置を例示したものに過ぎないことが理解されるべきである。例えば、上記の説明は人間の顔の認識に焦点を当てているが、本技術に熟練した者には、本発明の技術が、所与のオブジェクトが、その表示がデータベースに保存されている複数の既知のオブジェクトの1つと一致する、何らかの一般的な画像認識に適用可能なことが理解されるだろう。さらに、上記で説明された本発明の実施例は、例えば、入力ベクトルとコードベクトルの間の差を測定するために平均二乗誤差を利用しているが、他の実施例は、例えば、知覚的基準を組み込んだ知覚的計量を含む、他の誤差の計量を使うことができる。こうした知覚的計量は本技術に熟練した者には周知である。本技術に普通に熟練した者によって、本発明の精神と範囲から逸脱することなく、本発明の原理に従って無数の多様な装置が得られる。

【図面の簡単な説明】

【図1】既知の顔のためのデータが本発明の例示としての実施例に従って構成される、データベース発生プロセスの略図である。

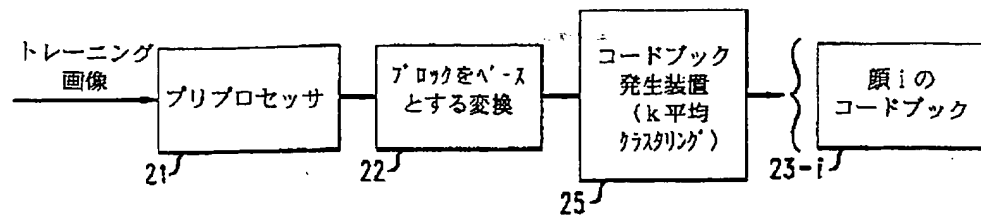
【図2】本発明の例示としての実施例による顔認識プロセスの略図である。

【図3】図1の例示としてのデータベース発生プロセスの例示としてのコードブック発生装置の流れ図である。

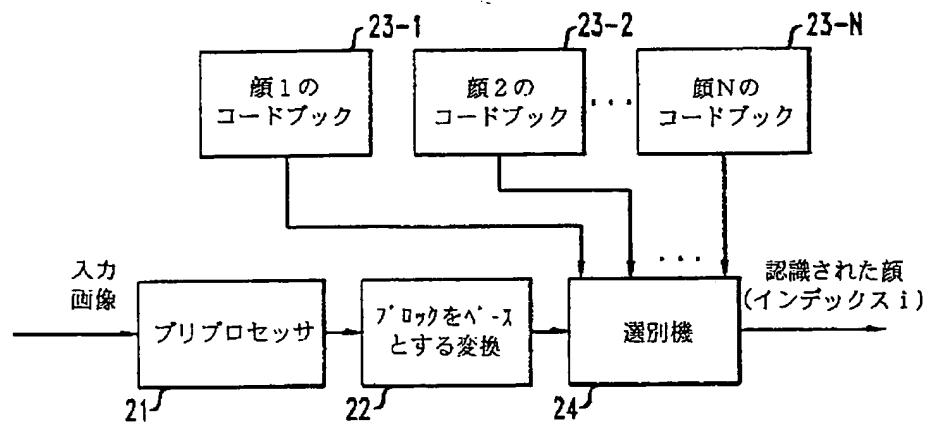
【図4】本発明の第1の例示としての実施例による図2の例示としての顔認識プロセスの選別機の流れ図である。

【図5】本発明の第2の例示としての実施例による図2の例示としての顔認識プロセスの選別機の流れ図である。

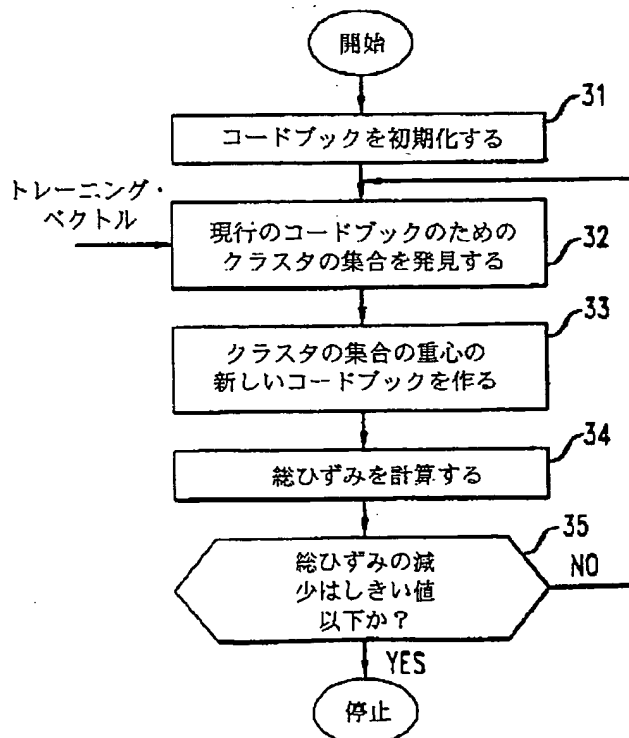
【図1】



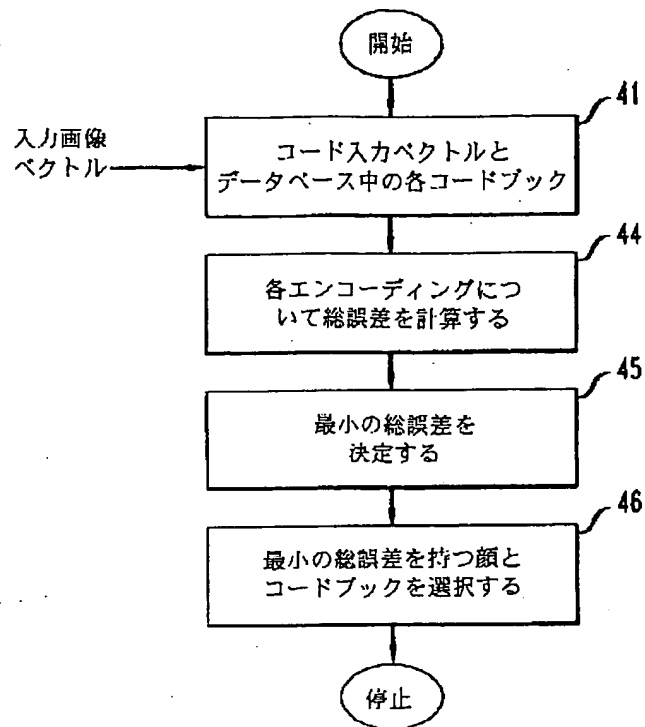
【図2】



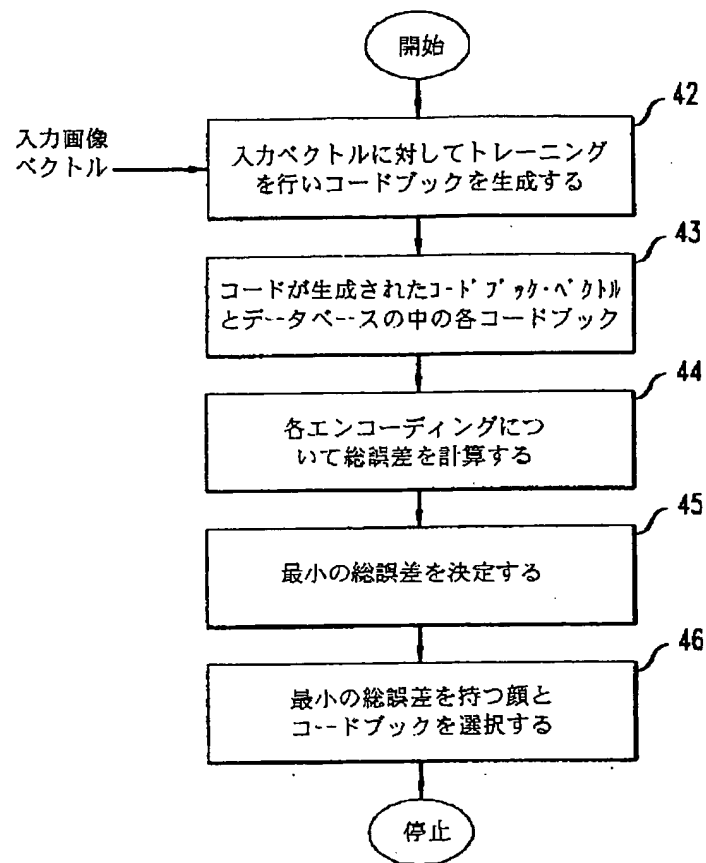
【図3】



【図4】



【図5】



フロントページの続き

(72)発明者 キシアオユ ツァン
アメリカ合衆国 08854 ニュージャージー
イ, ビスカッタウェイ, アpartment
ナンバー274, ロイヤル ドライブ 60